Keynote Address

Welcome and Introduction of Keynote Speaker Jay Hakes
Katherine K. Wallman
Office of Management and Budget

It is a special pleasure for me to welcome today's keynote speaker, Jay Hakes, who currently serves as the Director of the Jimmy Carter Presidential Library in Atlanta, Georgia. When Jay first told me that he was taking this position, I was somewhat surprised – envisioning a rather dry building filled with the records of the Carter presidency. But Jay advised me that much more is involved – and indeed, a few highlights he recently provided to me bear that out. For example, the archival materials at the library provide the foundation for an upcoming "American Experience" on PBS – a biography of Jimmy Carter that will run On November 11 and 12 – which I now plan to watch. The Museum associated with the library has just finished hosting the American Independence Road Trip with Norman Lear's copy of the Declaration of Independence. along with other great original documents from the Revolutionary War period. From September 27 to January 5, 2003, the Museum is hosting "American Originals," a collection of major original documents including the Louisiana Purchase, Edison's patent on the light bulb, the surrender documents from World War II, and the arrest warrant for Susan B. Anthony illegally voting. The exhibit also includes the Emancipation Proclamation, which has not come to the Southeast since 1949. I am confident that Jay could entertain us for the next hour ... and far more ... with vignettes from his current endeavors.

But why, you may be asking, did I suggest that Jay Hakes serve as the keynote speaker for our biennial Federal Committee on Statistical Methodology Seminar. Let me explain. As many of you know, Jay served as the presidentially appointed, Senate confirmed Administrator of the Energy Information Administration from 1993 to 2000. During that period, he was a principal spokesman on energy issues, briefing policy officials throughout the Federal government (and around the world), testifying frequently before congressional committees, and interacting regularly with news organizations. At the heart of Jay's efforts were a strong and steady commitment to making the products of EIA, and indeed the statistical system, more readily understandable by and accessible to the many policy makers and publics we serve. Thus, while Jay oversaw the development of EIA's award-winning web site, he also laid the foundation for further efforts. For example, that site has just been deemed "best site for tracking economic trends" by Time magazine. And, as a member of the Interagency Council on Statistical Policy, Jay strongly encouraged and supported the birth and maturation of FEDSTATS. Always, it seemed to me, Jay Hakes challenged his own agency, and his sister agencies, to be a bit more creative, a bit more assertive, and a bit more responsive to those who could benefit from the information we statisticians provide. His insights and his proposals always were respected – and acted favorably upon – by his colleagues around the agency heads table. We learned a great deal from Jay Hakes; we were fortunate that he was keen to serve as the head of EIA.

And so, it is with great personal and professional pleasure that I introduce Jay Hakes to challenge us as we strive to foster access to Federal statistics.

A Gift to the American People:
Victories and Challenges in Providing Web Access to Federal Statistics
Jay E. Hakes
Jimmy Carter Presidential Library

It's great to be with you today. I'd like to thank Kathy and Ed for inviting me. I'm delighted to be back with many friends and former colleagues. I am here for a reason. It's because I've always done what Kathy Wallman told me to do.

From the somewhat distant perspective of a presidential library, I'd like to repeat what I said before I left Washington. The technical competence and independent integrity of the statistical agencies contribute to the foundations of our democratic system. Whatever the future holds for our country, we need to not only retain these values, but encourage their continued development.

For those of you who are interested in what I do now, I suggested you watch the "American Experience" on PBS next Monday and Tuesday nights. They have produced a major new biography on President and Mrs. Carter. Most of the material came either directly or indirectly from the archives at the Carter Presidential Library.

I continue to be fascinated by our various national energy policies and the attempts of some to suggest their policy is the first of its kind. So I'm doing historical research on this issue in my spare time. Right now at the library we have a letter and sword sent by the King of Siam to the President of the United States. It part of a collection called "American Originals" that includes the Louisiana Purchase, the Emancipation Proclamation, and the arrest warrant for Susan B. Anthony illegally voting – all on loan from the National Archives here in Washington. When the sword was mailed from what is now modern Thailand, James Buchanan was president. By the time it arrived, Abraham Lincoln had taken office. In the letter, the King offered the President elephants to breed for national transportation needs. Lincoln responded that he wasn't sure that elephants would breed in our climate. Furthermore, he said we had committed to steam power on our rivers and rails. I think it's fair to say our national energy policies go at least as far back as Lincoln.

Well, today I've been asked to speak in a general way about access to federal statistics to kick off this conference. I can do so as a former producer of federal information at the Energy Information Administration and the Council on Statistical Policy, a current collector and sharer of presidential archives, and a frequent consumer of federal information of many kinds. Though in Atlanta, I'm only a click away from what you produce. I assure you I use it frequently.

The key word for today is "access."   This is a word that's achieved great cache in today's cyber world and in government circles. Maybe even too much cache:

♦ Access is the name, of course, of a popular Microsoft database.

♦ Adobe also has software named Access, which helps the blind and visually impaired read web documents.

♦ The State of Indiana calls its web site "AccessIndiana." In Arizona, it's "AccessArizona." In Idaho, it's "AccessIdaho." (I think you get the picture.)

♦ The web offers us access to wine, access to art, as well as, first and foremost, access to information.

♦ We can even find web sites that help us restrict access to unwanted information. A site called "NetNanny" can help if you have this problem. (I'm not making this up.)

♦ Access has been perhaps the most important word in the strategic plan of several federal agencies, including my former agency the Energy Information Administration and my current agency the National Archives and Records Administration, as they attempt to utilize electronic tools to accomplish their missions. If you look at the introduction to EIA's strategic written in 1994, it was all about access.

♦ More recently, I should also note that in September President Bush ordered the development of an interagency disability web site. The announcement promised the site would provide people with disabilities "access to a single point to go online for Government information and resources related to disabilities." Incidentally, the word "access" is used a couple of additional times in the announcement.

The federal statistical agencies have, of course, established strong web sites to encourage use of official data some time back now. The general site, FedStats, has always promised, "direct access to statistical data on topics of your choice."

Access to federal data involves more than just maintaining good web sites. But the change brought about by the web has been revolutionary. In fact, I find myself looking at the release of the Netscape browser as a fundamental turning point in the kinds of access we can and do provide today.

In my remarks, I'd like to talk about

♦ What access means for federal agencies,
♦ Some of the obstacles we've had to overcome to provide the access we have today,
♦ Some of the benefits we've gotten from our efforts, and
♦ Where we might best devote our future efforts.

It should come as no surprise that "access" can mean different things to different people.

I would make an important distinction between access that is grudging and passive and access that is expansive and active.

Grudging access can be associated with words like "bureaucratic" and "legalistic." At its worst, it's reflected in the attitude: "If this person has actually found out we have this stuff, I guess we might have to give it to them." Unfortunately, this kind of access is still the norm in a few places. (I won't name them, but I could.)

Expansive access, on the other hand, is associated with words like customer service, finding potential customers, and public education. Customers of government services are increasingly expecting this kind of access and increasingly they're getting it.

Easy access to government information is a hallmark of a democratic society. James Madison is often quoted for his comment: "popular Government, without popular information, or the means of acquiring it, is but a Prologue to a Farce or a Tragedy; or perhaps both." In today's complex world the range of issues and choices seems to have no end. Madison's sentiments point to the modern value of easy access to information, in a manner than goes well beyond the minimal requirements of the law.

The development of modern web sites began very recently, basically in the mid-1990's. So how did we get good statistical agency web sites so quickly?

Many in this room were involved in the early efforts. But some of you may have forgotten the obstacles we faced at the time:

♦ First, we didn't have a lot of young employees. So, if the stereotypes were correct, we shouldn't have been very web savvy.

♦ Second, there wasn't much, if any money appropriated for the specific purpose of developing web sites, so we could have easily justified inaction by a lack of resources.

♦ Third, our regular customers weren't demanding web-access in the mid-1990's, because they didn't have modems yet.

♦ Fourth, there were undoubtedly a few people in government who would have been very nervous about all this information going out if they had been alert enough to figure out what was going on.

♦ In addition, some employees were hesitant to move quickly. Some saw a focus on the web as a distraction from their "regular work." Others were wary of making information available to the masses in a way that experts wouldn't be there to "explain it." I even remember a discussion or two about the fairness of putting information up on the web, since it would be primarily the rich who would have the equipment to use it and would, as a result, gain superior access.

Another problem, at least at EIA, was a movement in congress in the mid-90's to have statistical agencies offset the cost of data collection by selling it. The House Budget Committee for two

consecutive years provided that EIA's appropriation would be cut in half on the assumption it could earn an equivalent amount from selling its products. (I think might have been meant as a compliment, albeit an unwelcome one.) The difficulty of private web sites earning money on the web from information they can copyright suggests the futility of trying to earn substantial revenues from electronic information that can't be copyrighted. One strategy would have been to restrict electronic access to increase the revenue potential of hard copies. This is an important point, because our brethren at the OECD and in some other countries went the direction of emphasizing sales of hard copies over electronic availability. Because they were forced to make revenue off their highly priced print publications, they couldn't give much information away free on the web. This policy, in effect, choked attempts to provide excellent customer service on the web.

Fortunately, the federal statistical agencies were able to overcome the obstacles I've described. Many of our employees were or became web savvy and found the challenge fun and exciting. They foresaw the need to provide good electronic access before customers were asking for it. I remember visiting the offices of sometimes-puzzled congressional staffers around 1995 demonstrating a data-rich CD-ROM at a time they had neither CD drives nor modems.

The web, of course, leapfrogged everything, and CDs never became the major player we thought they would. But when customers were ready to use web products, we were already providing them. If agencies had waited until they were asked to provide web sites, the necessary lead times for development would have put them well behind the curve.

Fortunately, it didn't cost a lot of money to develop products for the web. As a result, the financial constraints normal for government projects were minimized. The House Appropriations Committee ignored its direction from the Budget Committee to replace appropriations with sales revenues. The most obvious result of all these developments was very good web sites – content rich, relatively integrated, well tagged for search engines, with good navigability. I might also add that these government sites are much better than most business sites, despite some myths to the contrary.

What are the Benefits?
I've done a quick count of some benefits of developing good web sites. I'm sure there are many more, but I selected ten worth mentioning:

♦ Most obviously, people all over the country were able to access official data in a timely manner. Since the costs of to the taxpayers for this service was low, I call this a great gift to the American people. Previously, obtaining hard copies was at best slow and at worst virtually impossible. With our new web sites, we said: "A high school student in Altoona, PA, has much data available as a cabinet secretary had five years previously." But I also think of the reporter in California who working on a story at 5 Pacific Time, after offices in Washington are closed. Now the data are still available to them because the web sites are always open. I also remember staff from the National Economic Council at a conference in Buenos Aires tapping regularly into statistical agency web sites. I assure that in the days of

hard copies in the suitcase, these data would simply have been too bulky to travel and thus ignored.

♦ Second, web sites facilitated communication with those close by. A recent television ad portrays two business two construct a major business deal over the web, only to find they're located across the street. Have you seen it? A lot of times we don't recognize that people just down the hall at the Labor Department, at the Justice Department and elsewhere are using our information much more than when they had to use hard copies.

♦ Third, web access has helped statistical agencies get credit for the work they do. In the "old days," clever repackagers would sell federal data to clients for big price tags and often neglect to mention the source of the data. This gave the impression that the data would still be produced, even if the statistical agencies went away. Now the easiest way to get federal data is directly from federal web sites. You know what? I don't feel sorry for the repackagers, because the good ones will always be able to find good ways to add value and give credit where credit is due.

♦ Fourth, web sites give us important feedback from customers. Software is readily available track what parts of your site people are using and what parts get less usage. For example, EIA found that people liked summaries, which encouraged it to do more of them.

♦ Fifth, the Fedstats gateway to federal statistical sites made it easier for the layman to find federal statistical. To order a Ford automobile, you don't have to know the plant where it was manufactured. You should be able to find federal data without knowing which agency conducted the survey.

♦ Sixth, electronic dissemination helped agencies meet the requirements of the Government Performance and Results Act. With GPRA, we were expected to account for the outcomes of our activities. With web sites, we could produce good evidence of high and rising usage. We could also do sample surveys more easily on the value of the data to the users.

♦ Seventh, government experts were able to devote their attention to higher level matters. Before web sites, highly skilled analysts were spending time faxing data tables. It's hard to imagine now why that might be necessary. Experts can spend their time answering the hard questions that better utilize their special talents.

♦ Eighth, Fedstats helped fend off what I regarded as an unnecessary and probably counterproductive reorganization of federal statistical agencies. Periodically, people in the congress look at bringing at least some of the statistical agencies into one super statistical agency, a sort of Department of Homeland Statistics, so the right hand will know what the left hand is doing. When reorg gained some momentum in 1997, the work on Fedstats helped demonstrate to the Congress that reorganization was largely unnecessary. The statistical agencies were already talking to each other and had solved at least some of the problems that reorg was supposed to resolve.

♦ Ninth, it's fun to win awards. How many times does a government agency win awards? Yet many federal sites have won awards for there web sites. The most recent was in this week's Time magazine. EIA was picked as one of the 46 best web sites for business. In fact, it was of ten sites to receive a star. The description said: "For free research on a crucial industry, try this site from the Department of Energy, which forecasts future prices and trends for oil, gas and other petroleum products. In addition to statistical tables, the EIA produces clearly written reports that spell out in plain English what the numbers mean. It also features profiles of the energy sector in various countries and regions."

♦ Tenth and finally, the electronic world got us used to color graphics. Color became a standard feature because it was cheap and it made our data easier to understand. Color is a habit that's hard to break. I, for one, never what to go back to the days of one-color line graphs.

What about the future?

There are still many access issues for the future:

♦ We have so much information on our sites that *navigation* remains a major challenge. It is still possible for a fairly savvy user to not find something that's actually there. The battle to fully integrate sites hasn't been fully won. You might say people could call a help desk. But if they don't see something, it's hard to ask for it. You have to remember another thing. Men never ask for directions.

♦ There is also a strong argument for going back and putting up some of "the old stuff" on the web. This project would take some resources. However, now that we've shown what we can do with the "new stuff," we might be in a better position to argue for putting up available data and reports produced "pre-Netscape."

♦ It is in the public interest and our institutional interest that we be as visible and accessible as possible. We're doing very well with this, if Google searches are a good indicator. If you type in the subject matter covered by the federal statistical agencies with the words "statistics" or "data," in most cases, the relevant federal agency will show up first or second in a Google search. If you're not, that's a serious matter that needs to be addressed. Are your files well tagged? Are you meeting customer needs?

In my view, access was the leading goal of the 1990's. Electronic access helped us perform our missions better and helped build an expanded customer base for our efforts. We still need some of "our best and our brightest" working to make our information even more accessible.

In my opinion, however, the statistical agencies have won enough of battle to provide access to adopt a new dominant goal for the current decade. In my view, our greatest focus now should be (plastics? No, that was 1967) -- *credibility.*

Let me read from an August op ed piece in USA Today. It's written by a journalism professor at Duke about the news profession, but I think it applies to us as well:

> At the start of each college semester, I ask my students: "What is it that a news organization has to sell?" After all the predictable answers – news, facts, information – we arrive at the only one that matters: credibility. Unless news consumers are getting the news they need, presented with fairness and balance, they will find other ways to keep abreast of current events.

I think this is even more the case for statistical agencies.

Credibility is an area where statistical agencies have always done well and have a competitive edge over other providers of information. For instance, admitting when you make a mistake is one of the most important aspects of credibility. We've already done that and need to continue this practice.

There is a lot of inaccurate information being distributed over the web, even on sites that look respectable. For instance, I wanted to do some web research on the gift of a sword by the King of Siam to the United States – the story to which I referred earlier. Several sites say the King gave the sword in recognition of Lincoln's Emancipation Proclamation. The only problem with this assertion is that the King, as I mention earlier, mailed the sword while Buchanan. It was only because of the slowness of transportation in the 1860's that Lincoln was the recipient. There's a lot of information on the web that sounds good, but is in error.

Federal statistical agencies need to differentiate themselves from other sites. They should not mislead users who can now use the data for everything from policy speeches to investment decisions. This is an issue of both substance and image. With more people using our products, we are more vulnerable if any weaknesses in our systems are seen to be the cause of public misinformation.

Since the issue of credibility isn't the major thrust of my speech, I will only list a couple of the issues involved, and you'll have to invite me back if you want to discuss them in more detail:

♦ Response rates and quality. We all know that it's been getting harder to get the public to respond to requests for information. We have also seen reports that energy trading companies intentionally misreported data to private, but respected firms who compile and report data. The purpose of inaccurate reporting was to manipulate energy markets. We need to attack the problems of under and misreporting very aggressively. We cannot become resigned to these problems and begin treating them as necessary evils. We have to find solutions. I will be interested to see what suggestions come out of the conference in this regard.

♦ Timeliness. When major decisions hinge on official data, it is unfortunate when those data don't reflect current reality. Timeliness can be a threat to quality of data, if we're rushing out shoddy information. However, timeliness is an essential part of quality. As the computer has

been used to provide access, it can be used to reduce the time it takes to process data. If our processing times are not dramatically different than they were ten years ago, they are certainly taking much too long.

♦ I would add another issue that might be a bit sensitive. Cabinet-level departments sometimes see data as weapons to be used in behalf of advocacy of policy positions. In general, of course, this is fine. However, as the manufacturer of the bullets, the statistical agencies might be confused with the shooter of the bullets. As a result, statistical agencies need at times to retain a respectful distance from agencies in which they are housed. The credibility of the data is too valuable to risk.

If these comments look a lot like the agenda for this conference, I would point out that I completed my outline before I saw the titles for the panels. It sounds like there may be some consensus about the strategic issues we need to address. In the early 1990's, there was a lot of high-quality federal information for which the audience was too small. Now the audience has been greatly enlarged through web access. To keep and further enlarge that audience, we need to protect the quality standards we have and raise the bar for what quality means.